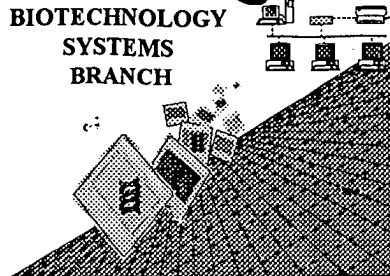


h/a

## RAW SEQUENCE LISTING ERROR REPORT

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



RECEIVED  
JUN 04 2001  
TECH CENTER 1600/2900

#4

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number:

09/595,326A

Source:

1638 RUSH

Date Processed by STIC:

5/31/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

### Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/595,326A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

RECEIVED  
JUN 04 2001  
TECH CENTER 1600/2900

- 1        Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2        Wrapped Aminos      The amino acid number/text at the end of each line "wrapped " down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3        Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4        Misaligned Amino Acid      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs  
Numbering      between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5        Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6        Variable Length      Sequence(s)        contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and  
indicate in the (ix) feature section that some may be missing.
- 7        PatentIn ver. 2.0 "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid  
sequence(s)       . Normally, PatentIn would automatically generate this section from the  
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section  
to the subsequent amino acid sequence. **This applies primarily to the mandatory <220>-<223>  
sections for Artificial or Unknown sequences.**
- 8        Skipped Sequences      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence:  
(OLD RULES)      **(2) INFORMATION FOR SEQ ID NO:X:**  
                                 (i) **SEQUENCE CHARACTERISTICS:**(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
                                 (xi) **SEQUENCE DESCRIPTION:SEQ ID NO:X:**  
                                 **This sequence is intentionally skipped**  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9        Skipped Sequences      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence.  
(NEW RULES)      **<210> sequence id number**  
                                 **<400> sequence id number**  
                                 **000**
- 10        Use of n's or Xaa's      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11        Use of "Artificial"      Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules.  
(NEW RULES)      Valid response is Artificial Sequence.
- 12        Use of <220>Feature      Sequence(s)        are missing the <220>Feature and associated headings  
(NEW RULES)      Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial Sequence" or "Unknown"  
                                 **Please explain source of genetic material in <220> to <223> section.**  
                                 **(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)**
- 13        PatentIn ver. 2.0 "bug"      **Please do not use "Copy to Disk" function of PatentIn version 2.0.** This causes a corrupted  
file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/595,326A

DATE: 05/31/2001  
TIME: 12:15:54

pp 1-5

Input Set : C:\PTO.txt

Output Set: C:\CRF3\05312001\I595326A.raw

Does Not Comply  
Corrected Diskette Needed

3 <110> APPLICANT: ALEXANDROV, Nickolai  
 4 TROUKHAN, Maxim  
 6 <120> TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding  
 Polypeptides Encoded

7 Thereby

9 &lt;130&gt; FILE REFERENCE: 2750-0942P

11 &lt;140&gt; CURRENT APPLICATION NUMBER: 09/595,326A

12 &lt;141&gt; CURRENT FILING DATE: 2000-06-16

14 &lt;160&gt; NUMBER OF SEQ ID NOS: 769

16 &lt;170&gt; SOFTWARE: PatentIn version 3.0

18 &lt;210&gt; SEQ ID NO: 1

19 &lt;211&gt; LENGTH: 5

20 &lt;212&gt; TYPE: PRT

21 &lt;213&gt; ORGANISM: Consensus Sequence

23 &lt;400&gt; SEQUENCE: 1

25 Leu Ile Val Met Thr

26 1 5

28 &lt;210&gt; SEQ ID NO: 2

29 &lt;211&gt; LENGTH: 5

30 &lt;212&gt; TYPE: PRT

31 &lt;213&gt; ORGANISM: Consensus Sequence

33 &lt;400&gt; SEQUENCE: 2

35 Leu Ile Val Met Phe

36 1 5

38 &lt;210&gt; SEQ ID NO: 3

39 &lt;211&gt; LENGTH: 5

40 &lt;212&gt; TYPE: PRT

41 &lt;213&gt; ORGANISM: Consensus Sequence

43 &lt;400&gt; SEQUENCE: 3

45 Gly Ala Thr Met Cys

46 1 5

48 &lt;210&gt; SEQ ID NO: 4

49 &lt;211&gt; LENGTH: 4

50 &lt;212&gt; TYPE: PRT

51 &lt;213&gt; ORGANISM: Consensus Sequence

53 &lt;400&gt; SEQUENCE: 4

55 Leu Ile Val Met

56 1

58 &lt;210&gt; SEQ ID NO: 5

59 &lt;211&gt; LENGTH: 4

60 &lt;212&gt; TYPE: PRT

61 &lt;213&gt; ORGANISM: Consensus Sequence

63 &lt;400&gt; SEQUENCE: 5

65 Leu Ile Phe Ala

66 1

68 &lt;210&gt; SEQ ID NO: 6

69 &lt;211&gt; LENGTH: 7

70 &lt;212&gt; TYPE: PRT

global error

invalid - per 1.823 of Sequence Rules, the only valid  
 <213> responses are: Unknown,  
 Artificial Sequence, or  
 Scientific name (Genus/species)  
 (one of the three)

(see circled portion of  
 item 12 on Error Summary  
 sheet)

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/595,326A

DATE: 05/31/2001

TIME: 12:15:54

Input Set : C:\PTO.txt

Output Set: C:\CRF3\05312001\I595326A.raw

71 <213> ORGANISM: Consensus Sequence  
73 <400> SEQUENCE: 6  
75 Leu Ile Val Met Phe Tyr Cys  
76 1 5  
78 <210> SEQ ID NO: 7  
79 <211> LENGTH: 11  
80 <212> TYPE: PRT  
81 <213> ORGANISM: Consensus Sequence  
83 <400> SEQUENCE: 7  
85 Ser Ala Pro Gly Leu Val Phe Tyr Lys Gln His  
86 1 5 10  
88 <210> SEQ ID NO: 8  
89 <211> LENGTH: 6  
90 <212> TYPE: PRT  
91 <213> ORGANISM: Consensus Sequence  
93 <400> SEQUENCE: 8  
95 Asp Glu Asn Gln Met Trp  
96 1 5  
98 <210> SEQ ID NO: 9  
99 <211> LENGTH: 12  
100 <212> TYPE: PRT  
101 <213> ORGANISM: Consensus Sequence  
103 <400> SEQUENCE: 9  
105 Lys Arg Gln Ala Ser Pro Cys Leu Ile Met Phe Trp  
106 1 5 10  
108 <210> SEQ ID NO: 10  
109 <211> LENGTH: 9  
110 <212> TYPE: PRT  
111 <213> ORGANISM: Consensus Sequence  
113 <400> SEQUENCE: 10  
115 Lys Arg Asn Gln Ser Thr Ala Val Met  
116 1 5  
118 <210> SEQ ID NO: 11  
119 <211> LENGTH: 7  
120 <212> TYPE: PRT  
121 <213> ORGANISM: Consensus Sequence  
123 <400> SEQUENCE: 11  
125 Lys Arg Ala Cys Leu Val Met  
126 1 5  
128 <210> SEQ ID NO: 12  
129 <211> LENGTH: 9  
130 <212> TYPE: PRT  
131 <213> ORGANISM: Consensus Sequence  
133 <400> SEQUENCE: 12  
135 Leu Ile Val Met Phe Tyr Pro Ala Asn  
136 1 5  
138 <210> SEQ ID NO: 13  
139 <211> LENGTH: 6  
140 <212> TYPE: PRT

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/595,326A

DATE: 05/31/2001

TIME: 12:15:54

Input Set : C:\PTO.txt

Output Set: C:\CRF3\05312001\I595326A.raw

141 <213> ORGANISM: Consensus Sequence  
143 <400> SEQUENCE: 13  
145 Leu Ile Val Met Phe Trp  
146 1 5  
148 <210> SEQ ID NO: 14  
149 <211> LENGTH: 8  
150 <212> TYPE: PRT  
151 <213> ORGANISM: Consensus Sequence  
153 <400> SEQUENCE: 14  
155 Ser Ala Gly Cys Leu Ile Val Pro  
156 1 5  
158 <210> SEQ ID NO: 15  
159 <211> LENGTH: 5  
160 <212> TYPE: PRT  
161 <213> ORGANISM: Consensus Sequence  
163 <400> SEQUENCE: 15  
165 Phe Tyr Trp His Pro  
166 1 5  
168 <210> SEQ ID NO: 16  
169 <211> LENGTH: 4  
170 <212> TYPE: PRT  
171 <213> ORGANISM: Consensus Sequence  
173 <400> SEQUENCE: 16  
175 Lys Arg His Pro  
176 1  
178 <210> SEQ ID NO: 17  
179 <211> LENGTH: 10  
180 <212> TYPE: PRT  
181 <213> ORGANISM: Consensus Sequence  
183 <400> SEQUENCE: 17  
185 Leu Ile Val Met Phe Tyr Trp Ser Thr Ala  
186 1 5 10  
188 <210> SEQ ID NO: 18  
189 <211> LENGTH: 6  
190 <212> TYPE: PRT  
191 <213> ORGANISM: Consensus Sequence  
193 <400> SEQUENCE: 18  
195 Leu Ile Val Met Phe Tyr  
196 1 5  
198 <210> SEQ ID NO: 19  
199 <211> LENGTH: 4  
200 <212> TYPE: PRT  
201 <213> ORGANISM: Consensus Sequence  
203 <400> SEQUENCE: 19  
205 Gly Ser Thr Ala  
206 1  
208 <210> SEQ ID NO: 20  
209 <211> LENGTH: 4  
210 <212> TYPE: PRT

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/595,326A

DATE: 05/31/2001

TIME: 12:15:54

Input Set : C:\PTO.txt

Output Set: C:\CRF3\05312001\I595326A.raw

211 &lt;213&gt; ORGANISM: Consensus Sequence

213 &lt;400&gt; SEQUENCE: 20

215 Ser Thr Ala Gly

216 1

218 &lt;210&gt; SEQ ID NO: 21

219 &lt;211&gt; LENGTH: 4

220 &lt;212&gt; TYPE: PRT

221 &lt;213&gt; ORGANISM: Consensus Sequence

223 &lt;400&gt; SEQUENCE: 21

225 Ser Thr Glu Ile

226 1

228 &lt;210&gt; SEQ ID NO: 22

229 &lt;211&gt; LENGTH: 7

230 &lt;212&gt; TYPE: PRT

231 &lt;213&gt; ORGANISM: Consensus Sequence

233 &lt;400&gt; SEQUENCE: 22

235 Pro Ala Ser Leu Ile Val Met

236 1 5

238 &lt;210&gt; SEQ ID NO: 23

239 &lt;211&gt; LENGTH: 4

240 &lt;212&gt; TYPE: PRT

241 &lt;213&gt; ORGANISM: Consensus Sequence

243 &lt;400&gt; SEQUENCE: 23

245 Ser Ala Lys Arg

246 1

248 &lt;210&gt; SEQ ID NO: 24

249 &lt;211&gt; LENGTH: 5

250 &lt;212&gt; TYPE: PRT

251 &lt;213&gt; ORGANISM: Consensus Sequence

253 &lt;400&gt; SEQUENCE: 24

255 Ser Thr Ala Gly Asn

256 1 5

258 &lt;210&gt; SEQ ID NO: 25

259 &lt;211&gt; LENGTH: 4

260 &lt;212&gt; TYPE: PRT

261 &lt;213&gt; ORGANISM: Consensus Sequence

263 &lt;400&gt; SEQUENCE: 25

265 Ser Ala Gly Val

266 1

268 &lt;210&gt; SEQ ID NO: 26

269 &lt;211&gt; LENGTH: 7

270 &lt;212&gt; TYPE: PRT

271 &lt;213&gt; ORGANISM: Consensus Sequence

273 &lt;400&gt; SEQUENCE: 26

275 Leu Ile Val Met Phe Tyr Trp

276 1 5

278 &lt;210&gt; SEQ ID NO: 27

279 &lt;211&gt; LENGTH: 7

280 &lt;212&gt; TYPE: PRT

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/595,326A

DATE: 05/31/2001

TIME: 12:15:54

Input Set : C:\PTO.txt

Output Set: C:\CRF3\05312001\I595326A.raw

281 &lt;213&gt; ORGANISM: Consensus Sequence

283 &lt;400&gt; SEQUENCE: 27

285 Gly Ser Ala Cys Ile Val Met

286 1 5

288 &lt;210&gt; SEQ ID NO: 28

289 &lt;211&gt; LENGTH: 4

290 &lt;212&gt; TYPE: PRT

291 &lt;213&gt; ORGANISM: Consensus Sequence

293 &lt;400&gt; SEQUENCE: 28

295 Gly Thr Ile Val

296 1

298 &lt;210&gt; SEQ ID NO: 29

299 &lt;211&gt; LENGTH: 6

300 &lt;212&gt; TYPE: PRT

301 &lt;213&gt; ORGANISM: Consensus Sequence

303 &lt;400&gt; SEQUENCE: 29

305 Gly Ser Thr Ala Asn Ile

306 1 5

308 &lt;210&gt; SEQ ID NO: 30

309 &lt;211&gt; LENGTH: 5

310 &lt;212&gt; TYPE: PRT

311 &lt;213&gt; ORGANISM: Consensus Sequence

313 &lt;400&gt; SEQUENCE: 30

315 Leu Ile Val Met Ala

316 1 5

318 &lt;210&gt; SEQ ID NO: 31

319 &lt;211&gt; LENGTH: 6

320 &lt;212&gt; TYPE: PRT

321 &lt;213&gt; ORGANISM: Consensus Sequence

323 &lt;400&gt; SEQUENCE: 31

325 Leu Ile Val Trp Pro Gln

326 1 5

328 &lt;210&gt; SEQ ID NO: 32

329 &lt;211&gt; LENGTH: 5

330 &lt;212&gt; TYPE: PRT

331 &lt;213&gt; ORGANISM: Consensus Sequence

333 &lt;400&gt; SEQUENCE: 32

335 Gly Ser Thr Ala Met

336 1 5

338 &lt;210&gt; SEQ ID NO: 33

339 &lt;211&gt; LENGTH: 6

340 &lt;212&gt; TYPE: PRT

341 &lt;213&gt; ORGANISM: Consensus Sequence

343 &lt;400&gt; SEQUENCE: 33

345 Leu Ile Met Pro Thr Ala

346 1 5

348 &lt;210&gt; SEQ ID NO: 34

349 &lt;211&gt; LENGTH: 4

350 &lt;212&gt; TYPE: PRT

*Please correct this error  
in subsequent sequences.*

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/595,326A

DATE: 05/31/2001

TIME: 12:15:55

Input Set : C:\PTO.txt

Output Set: C:\CRF3\05312001\I595326A.raw